

Selection of Reproductive Health End Points for Environmental Risk Assessment

by David A. Savitz* and Sioban D. Harlow*

In addition to the challenges inherent in environmental health risk assessment, the study of reproductive health requires thorough consideration of the very definition of reproductive risk. Researchers have yet to determine which end points need to be considered to comprehensively evaluate a community's reproductive health. Several scientific issues should be considered in the selection of end points: the severity of the outcomes, with a trade-off between clinical severity and statistical or biological sensitivity; the relative sensitivity of different outcomes to environmental agents; the interrelationship among adverse outcomes; the baseline frequency of the adverse outcome; evidence from reproductive toxicology; and specificity of reproductive effects from the environmental agent. Simultaneously, practical concerns should be addressed: frequency of occurrence of an event and consequent statistical power to evaluate changes; frequency of prerequisites (e.g., pregnancy) that are necessary to be at risk; time and money resource requirements for measuring the outcome; amenability of the end point to retrospective measurement; and burden of measurement on the population being studied. In this article, we discuss these scientific and practical considerations and recommend that reproductive risk assessment include measures of fecundability (menstrual function, time to pregnancy), fetal loss (clinically recognized miscarriage), and infant health (birth weight, gestational age). Additional methodological research is needed to refine the array of reproductive health measures that need to be examined as a consequence of environmental exposures.

Introduction

In contrast to cancer or cardiovascular disease, reproduction is not a disease entity. Although this statement seems both obvious and simplistic, the fact that reproduction refers to a biologic and social process as opposed to a specific pathologic entity has profound implications for the scientist or policy maker interested in reproductive risk assessment. In addition to the familiar methodological challenges that confront investigators of environmental influences on chronic disease, such as exposure ascertainment and control of confounding variables, reproductive epidemiologists are still grappling with the most fundamental question: What constitutes reproductive risk and how do we measure it?

Successful reproduction involves several components: the biologic processes of the male and female that yield viable sperm and a viable egg; the social and biologic interactions between the male and the female that yield a viable conceptus; the biologic processes within the mother and fetus that result in appropriate growth and development of the fetus; and the processes that ensure successful parturition. Disruption of any of these components may interfere with reproduction, but the manifestation of a disruption depends on the affected component and the nature of the insult. Multiple effects may result from interference with a specific biological process; conversely, each manifestation of reproductive health has several potential etiologic pathways.

For example, a woman whose endocrine function is disturbed may not ovulate, thus precluding pregnancy. The failure to ovulate may be manifested as an alteration in her menstrual cycles, or it may not become apparent until she tries to conceive and fails. On the other hand, if the alteration in endocrine function interferes with fetal development as opposed to ovulation, the woman will become pregnant, but the fetus may die and the mother may miscarry, or the fetus may survive with a serious birth defect.

Although information about environmental influences on reproduction is quite limited, available data suggest that essentially every component of the reproductive process is susceptible to some exogenous agent. Ovarian function is disrupted by exposure to chemotherapeutic drugs (1), spermatogenesis is disrupted by exposure to dibromochloropropane (2), smoking has been associated with delayed conception (3) and low birth weight (4), and developmental disabilities have occurred among children exposed to methyl mercury *in utero* (5).

Despite this multifaceted vulnerability to environmental insults (6,7), researchers concerned with a potentially toxic exposure often fail to consider multiple reproductive end points. Most often, only one specific end point is studied, and methodological efforts focus on how to better assess that individual outcome. For example, considerable attention is currently being addressed to how best to evaluate pregnancy loss. Pregnancy loss is clearly an important event for which etiologic factors need to be identified. The goal of environmental risk assessment, however, is not simply to elucidate the etiology of a particular adverse outcome, but to assess the overall reproductive health consequences of exposure to environmental agents.

*Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, NC 27599.

Address reprint requests to D. A. Savitz, Department of Epidemiology, CB #7400, School of Public Health, University of North Carolina, Chapel Hill, NC 27599.

The disease-based epidemiologic model leads to the evaluation of various individual end points, such as conception, fertility, or child health. Reproductive risk encompasses this entire spectrum. Consequently, investigation of reproductive risk and the accompanying methodological development must also address this broader concept. In parallel with current efforts to better evaluate specific outcomes, some thought and effort should therefore be devoted to integration of information across the spectrum of reproductive health.

This paper raises several methodological issues arising from the breadth and complexity of reproductive function. A simple, direct question underlies the discussion: Given the exposure of a community to a potentially toxic agent, how should we assess the reproductive health consequences? Specifically, which reproductive end points should we evaluate?

Selecting Reproductive End Points for Evaluation

Unfortunately, it must be conceded at the outset that our current level of knowledge precludes the formulation of explicit guidelines that will ensure the right choice of reproductive end points. More comprehensive consideration of the components of reproductive health is necessary and desirable for public health protection, yet an exhaustive evaluation that would ensure the absence of any health hazard is not feasible due to practical constraints (time, finances, population size). Recognizing that negative results are always subject to the challenge that the most sensitive health measure was omitted, more complete discussion of the reasoning behind the selection of reproductive end points would be instructive, as would post hoc examinations of the accuracy of the assumptions which guided the initial selection. Our limited knowledge should encourage innovation in developing new markers of reproductive health, such as the recently proposed indicator of time to pregnancy (8), while simultaneously evaluating traditional indicators such as fetal loss (9) and birth weight (10).

Scientific Considerations in Selecting Reproductive End Points

What is the Role of Clinical Severity in Selecting Reproductive End Points?

The clinical significance of a reproductive health event refers to the severity of the outcome for the physical or mental health of the mother or child. Events such as very low birth weight (< 1500 g), stillbirth, or major birth defects are considered severe due to their life-threatening nature or need for medical intervention, whereas delayed conception, subclinical miscarriage, or minor birth anomalies are not. All other considerations being equal, the most important outcomes to study are those with the greatest health impact. For a variety of reasons, however, other considerations are often not equal.

Rarity often precludes the study of the most serious reproductive events (e.g., specific major birth defects), especially when evaluating small communities. In some instances, outcomes that are measured on a continuous scale and dichotomized for clinical purposes may yield more information if evaluated as continuous

measures. For example, birth weight is truncated at 2500 (low birth weight) or 1500 (very low birth weight), and time to pregnancy is dichotomized at 1 year to approximate the clinical definition of infertility (11). In both these situations, only events in the tails of the distributions are analyzed as clinically significant adverse outcomes. Shifts within the dominant part of the distribution are ignored.

An argument can be made, however, for studying the entire distribution of such outcomes. First, shifts in the overall distribution rather than the proportion falling into the highly deviant tail may actually be the most useful indicator of the population's health. The magnitude of risk from shifts within the normal range of time to pregnancy, for example, may be inconsequential for the individual, but important on the population level. Statistical power considerations also favor studying shifts in a continuous measure rather than changes in the proportion of the population that is substantially deviant. The underlying assumption of this approach is that relatively modest shifts of the entire distribution of birth weight, for example, will move some otherwise healthy infants into the more clinically dangerous low birth weight zone even when the population is not large enough to examine an increased risk of this more severe outcome directly. This assumption warrants explicit evaluation.

Another argument for studying less severe outcomes is that they often more directly reflect the biological process of concern. For example, infertility is accepted as a clinically significant condition that can result from a disturbance in numerous biological mechanisms. By careful evaluation of such clinically insignificant events as a longer time to pregnancy, perturbations in the menstrual cycle, or altered sperm motility, the biological processes underlying infertility may be better understood and studied with greater precision. In fact, the availability of such subclinical windows into the underlying biological process is an important consideration in reproductive epidemiology. In much the same way that cytogenetic markers may reflect events related to carcinogenesis (12), such indicators of reproductive function may be markers of more severe pathology and potentially allow for earlier detection of adverse effects. However, given our limited knowledge about these functional yet clinically insignificant measures, investigators are obligated to demonstrate empirical as well as theoretical links to the more severe outcomes of ultimate public health concern.

Is There Consistency in Which Reproductive End Points Are Vulnerable to Environmental Agents?

The most fortuitous scenario for evaluating reproductive risk would be to have a single reproductive end point that was consistently found to be most readily perturbed by environmental exposures. Studies could then focus on this single end point, providing reassurance of safety if negative results were obtained, and proceeding to examine other end points only if an adverse effect were found.

The varied pathways through which reproductive function can be disturbed make such consistency unlikely. For example, the role of dibromochloropropane in producing testicular atrophy (13) has no utility for predicting the potential sensitivity of female reproductive function or fetal development to this agent. Although that it is possible that a pattern will emerge for some

classes of agents (e.g., solvents), investigators currently need to examine multiple end points.

The temporal relationship linking exposure to health outcomes would also be expected to vary across reproductive events. As recommended by Rothman (14), specific hypotheses should be made regarding the duration of exposure required for induction of disease and the time after induction during which the disease will not be apparent (latency). The traditional chronic disease model of multiyear induction and latent periods is probably not applicable to most reproductive health outcomes, so thought and flexibility are called for in evaluating temporal relationships. Prepregnancy exposures might be expected to affect fertility (through various pathways in the male and female). First trimester exposures would potentially affect fetal viability or produce birth defects. Second or third trimester exposures might affect complications of pregnancy, fetal growth, or timing of delivery. The generally brief time periods over which exposure can affect reproductive outcomes calls for unusual precision in measuring the timing of exposure but has the advantage of allowing examination of changes in health relative to changes in exposure over relatively short intervals (several years).

Does an Adverse Effect on One End Point Predict Adverse Effects on Others?

Given the need to evaluate more than one end point, an important consideration in selection of reproductive health indices is the potential redundancy of information. If outcomes are highly correlated at the individual level, as altered menstrual patterns and altered ovarian function are, measurement of both outcomes may be unnecessary. Expenditure of resources on the more challenging and expensive hormonal measurement of ovulation would be unjustified if more readily available data on menstrual bleeding patterns would address potential risk. If the correlation is at a community rather than individual level (e.g., if exposure has a similar effect on the risk of both clinically recognized miscarriages and subclinical miscarriages), evaluation of the less costly and less invasive marker is preferred.

The potential for competition among end points that occur along a continuum must also be considered. An extremely hazardous exposure that increases the frequency of a common early event may actually prevent a later adverse outcome (15). The classic example is the selective loss of fetuses who would, if they survived, have had congenital anomalies at birth (16). If an agent effectively causes early fetal loss, the future stillbirths or birth defect cases may, at least in theory, be prevented.

A complete understanding of the biological processes underlying the different reproductive end points could, ideally, predict the interrelationships among reproductive health measures and lead to a parsimonious selection of the most convenient indicators. Chromosomal damage to germ cells, for example, could potentially result in an array of adverse outcomes including infertility, fetal loss, birth defects, and childhood or adult disease (17). Stein et al. (16) have argued that spontaneous abortions constitute a preferred indicator of genetic damage as compared to the study of birth defects because of their greater frequency of occurrence. On the other hand, measurement of outcomes with distinct biological pathways would not yield redundant information. Exposures that suppress ovulation, for example, would not

necessarily be expected to have any correlation with influences on infant birth weight.

Our understanding of biological mechanisms is so limited, however, that more empirical evidence regarding the interrelationships among measures is needed in order to determine which items to include in an efficient assessment battery. For example, better data on the empirical relationships among subclinical pregnancy loss, early and late spontaneous abortion, and stillbirth would be very instructive regarding how much effort is actually needed to effectively study fetal loss. Generation of data that will enable development of efficient, comprehensive assessment will require simultaneous monitoring of several different outcomes at the individual and community level. Some considerations in efficiently selecting such end points are noted below.

Does Baseline Failure Rate Predict Environmental Sensitivity?

The spectrum of fallibility for reproductive events ranges from the very rare (e.g., ectopic pregnancy, specific birth defects) to the very common (e.g., early pregnancy loss). The most error-prone events, those with the highest natural incidence rates, might be the most sensitive to disturbance by environmental agents. The rationale for this proposition is intuitive rather than empirical. If the reserve capacity for a given reproductive function is so limited that natural failure occurs with a relatively high frequency, then each increment in stress to the system would be expected to further increase the number of adverse outcomes. Conversely, a process with substantial ability to withstand insult or recover may be better able to tolerate environmental stressors without observable effects. However, very common events such as early fetal loss may only reflect natural fallibility and may not be responsive to environmental influences. Empirical research is needed to determine which of these propositions is true.

The concept of baseline failure rate may also be applied to subpopulations. In principle, populations consist of immune (disease-free regardless of exposure), doomed (diseased regardless of exposure), and vulnerable (susceptible to exposure-induced disease) individuals (18). Populations with the highest proportion of vulnerable individuals would be expected to yield the strongest exposure-disease associations. The baseline frequency of adverse outcomes within a population may be a marker of such vulnerability. For many adverse reproductive outcomes such as spontaneous abortion or infertility, older women are at higher risk. Thus, it might be expected that older women, who naturally have a greater tendency for reproductive problems, would be most sensitive to environmental insults.

Alternatively, it might also be argued that high-risk persons have, by definition, other etiologic factors present that would make the isolation of an environmental contribution more difficult. Furthermore, if the increment in risk from exposure is additive, then a low baseline frequency would facilitate detection of an effect. Rothman and Poole (19) have suggested that in many circumstances, low-risk populations are most likely to manifest detectable effects of exposure. In the few studies that have examined the effect of exogenous agents among women with differing histories of fetal loss (20,21), the effects of exposure on high-risk women, defined as having a history of fetal loss, were

markedly less than the effects on women without such a history. Further consideration of exposure effects across the spectrum of baseline risk is warranted.

Can Toxicological Data Indicate the Critical End Points to Study?

Previous comments have focused on inherent characteristics of the reproductive outcome as the basis for selecting end points. Characteristics of the exposure should also influence this decision. Ideally, animal experiments with the environmental agents of interest would define the most relevant reproductive health end points to evaluate after human exposure to those agents. The complete lack of toxicological data on most agents is the primary limiting factor. In addition, strategies for making extrapolations from animals to humans on a qualitative, let alone quantitative, basis are at an early stage of development. Since each species, including humans, seems to have unique reproductive processes, the potential for extrapolation is uncertain (22). When direct extrapolation of reproductive end points from animals to humans is clearly not possible, available toxicological data may still offer some general guidance regarding the most likely target host (male, female, or fetus).

Is There Specificity for the Effect of an Environmental Agent?

With few exceptions, environmental reproductive toxins do not seem to have unique effects. Polychlorinated biphenyls have been linked to a pattern of discoloration at birth (23), and some medications are associated with specific birth defects (24), but most suggested hazards affect common reproductive outcomes such as spontaneous abortion and low birth weight. Although selection of common outcomes improves the statistical power to identify associations, the advantage is counterbalanced by the difficulty of identifying etiologic associations for end points that are adversely affected by many different agents.

Frequently, the outcome of interest is strongly predicted by lifestyle factors (e.g., cigarette smoking, alcohol consumption) or reproductive history. The distribution of environmental agents may in some cases introduce confounder-exposure associations as well. The potential link of hazardous exposures in the home or workplace with lower socioeconomic status, for example, could result in confounding by other correlates of social class such as cigarette smoking (25,26) or inadequate prenatal care. The challenge of isolating effects of environmental agents from lifestyle factors is commonly exacerbated by an interest in subtle effects in the presence of potentially strong confounding. Therefore, one important direction for research is to identify subgroups of reproductive events that are more directly and specifically affected by environmental exposures. For example, Kline and Stein (27) divided spontaneous abortions into those that are karyotypically normal versus those that are karyotypically abnormal in an attempt to discover clearer etiologic associations. Determinants of preterm delivery are similarly worth distinguishing from determinants of small for gestational age (4).

Practical Considerations in Selecting Reproductive End Points

The preceding discussion addressed scientific concerns in the selection of end points for evaluating environmental reproductive risks. This section addresses practical considerations in making such selections based on the feasibility of measurement.

Which End Points Are Sufficiently Common to Study in Small Populations?

The ability to precisely measure the frequency of an event and alterations in that frequency is a function of the spontaneous rate of occurrence of the event. Since exposures in the community or workplace often occur in relatively small populations, limitations imposed by modest study size must be recognized in the initial selection of end points. Obviously, more common events such as spontaneous abortion are favored over rarer events such as specific birth defects because small increments in risk can be more precisely measured for common end points. For example, to achieve 80% power to detect a doubling of risk in a cohort study with equal numbers of exposed and unexposed subjects would require approximately 125 pregnancies per group to examine fetal loss and over 700 infants per group to study all major birth defects (aggregated) (28). Furthermore, the public health importance of modest increases in a common event would be of greater significance than a comparable increase in a rare event.

Rarity is of much less concern when an exposure has highly specific effects. Specific birth defects constitute the most dramatic example of such specificity, in which a few cases (or even a single case) may identify an environmental hazard. Actually, such small clusters constitute marked multiplications in baseline risk, or, at the extreme, an infinite multiplication of risk of an otherwise nonexistent event. The attraction of identifying such clear cause-and-effect relationships is tempered by the relative infrequency with which such specificity has been demonstrated.

What Behaviors Are Required To Be at Risk for Specific End Points?

The ease with which an end point can be studied depends in part on the attributes or experiences that are required to contribute information about a possible hazard. For instance, evaluation of fertility requires a population of couples who are engaging in unprotected intercourse, while evaluation of birth outcomes requires a population of pregnant women. Operationally, these required attributes are defined by the units (persons, pregnancies, births) that appear in the denominator of rates or proportions.

Because available study size is greatly influenced by such requirements, those end points that do not require pregnancy or an attempt to become pregnant are favored. Nearly all reproductive-age males and females can contribute information on semen characteristics and menstrual function, regardless of marital or fertility status. The unit at risk for miscarriage is a pregnancy. The requirement of pregnancy substantially reduces the size of the informative population. More demanding and restrictive is

the requirement for a planned pregnancy, since only about half of all pregnancies are planned (29). Outcomes such as early fetal loss, time to pregnancy, and diagnosed infertility are only detectable among planned pregnancies. A slightly less restricted set of events consists of live births, the risk unit for preterm delivery or birth defects. For studies in which outcomes are passively monitored through medical diagnoses, an additional implicit requirement defining the sample population is the seeking of medical care (30). Estimates of available study size must therefore focus not on the total number of people in a community but on the number of informative events.

Which End Points Can Be Monitored with Limited Resources?

Financial constraints affect the feasibility of conducting studies as well as their geographic and temporal scope. The most desirable end points, based on financial considerations, are those that are reported in vital records. Registries can offer a unique resource due to their comprehensiveness, accessibility on an aggregate population basis, and historical availability. However, they can only be used to monitor certain outcomes, such as stillbirth, birth weight, gestational age, and some birth defects.

Ascertaining reproductive outcomes through medical records is more challenging. Hospital records are useful in identifying certain complications of pregnancy, such as pre-eclampsia and birth defects. Records from private clinics or physicians are much more difficult to obtain, though outcomes such as treated spontaneous abortions (9) and infertility (11) can be studied in this manner.

Self-reported reproductive outcomes are typically the most costly to identify since they require direct communication with individuals. No alternatives exist, however, for a number of outcomes, such as earlier spontaneous abortions and menstrual function. Finally, studies requiring medical examination or laboratory assays including semen analysis, hormone assays, or monitoring of subclinical events are the most expensive to conduct.

Can the End Point Be Accurately Reconstructed from the Past?

Since researchers are often interested in environmental exposures that occurred in the past, the amenability of an end point to accurate and complete historical reconstruction is often of interest. As noted previously, vital and medical records are typically available historically but are restricted in scope and often in quality (31,32). On the other hand, self-report of past events is dependent on the subjective importance of the event to the individual. Whereas live births are well recalled, spontaneous abortions are incompletely remembered (33). Spontaneous abortions of later gestational age are more completely recalled than those of earlier gestational age (33). For routine events such as menstrual abnormalities or early recognized pregnancy loss, prospective monitoring may be required (34). Subclinical events requiring laboratory identification are also not amenable to historical reconstruction, requiring true prospective ascertainment.

What Will People Tolerate in Monitoring Reproductive Health?

In addition to considerations of study size and cost, which affect feasibility from the investigator's perspective, the community's tolerance for participating in reproductive studies must also be addressed. Reproductive behavior is a sensitive topic. This sensitivity comes not only from the obvious sexual aspect of these behaviors, but also from the stigma associated with pregnancies to unmarried women and the social and moral concerns regarding induced abortion. Privacy concerns often surround the decision to become pregnant, especially with respect to employment.

Another dimension of the issue of respondent tolerance is the potential burden imposed by participation in protocols with intensive collection requirements such as the prospective monitoring of biological specimens for ovarian function or early pregnancy loss. The need to design studies that monitor multiple end points increases the potential for unacceptable respondent burden. Protocols that include appropriate encouragement and/or financial incentives can successfully retain participation in such demanding ventures, but the financial and human resources should be commensurate with the yield of information.

Suggested End Points for Environmental Reproductive Risk Assessment

In spite of the emphasis in this paper on articulating multiple considerations in selecting end points, the overriding importance of plausible biological mechanisms should not be lost. If a scientific basis exists for anticipating a specific consequence of an exposure, those specific end points should be the principal focus of any research. Logistical considerations in that selection process are secondary to the scientific principles. Given the frequent absence of guidance from toxicology or prior epidemiologic studies, however, we would suggest that environmental risk assessments include an array of relatively common, easily measured reproductive end points. Selection of an array of end points is preferable to focusing on one or two arbitrarily chosen measures, such as specific birth defects. The array should be chosen to include diverse processes related to reproductive health as well as different time courses of exposure and response. Based on these criteria, we recommend evaluation of fecundability, fetal survival, and infant health.

Fecundability. Important measures of reproductive function related to conception and early pregnancy viability would include markers of menstrual function (menstrual cycle length) as indicators of female endocrine function and time to pregnancy as a summary indicator of the complete array of processes required to conceive. Both are readily self-reported and do not require invasive technology, although prospective data collection is necessary for monitoring menstrual function (34). The relevant exposure would in part be contemporaneous with the manifestation of these outcomes.

Fetal Survival. Spontaneous abortion has been proven to be vulnerable to environmental agents. Practical considerations favor measurement of self-reported or medically treated spontaneous abortions (covering the period starting about 8–10 weeks into pregnancy), since measurement at earlier gestational ages requires expensive and demanding procedures (9).

Pregnancy loss is thought to result primarily from exposures during early pregnancy, which also is the vulnerable period for birth defects. Pregnancy loss, however, is a much more frequent occurrence.

Infant Health. The traditional and readily ascertained measures of infant health much as gestational age (prematurity) and birth weight warrant consideration. These outcomes are readily and relatively accurately ascertained from vital records, medical records, or self-report and have been shown to respond to a variety of exogenous influences (4). Furthermore, the later part of pregnancy (third trimester) is thought to be the most critical time period for an impact of potentially hazardous exposures.

Need for Methodological Evaluation

The current level of knowledge, in combination with societal demand and public health prudence, supports the conduct of epidemiologic research to evaluate the impact of environmental agents on reproductive health. Nonetheless, the uncertainties regarding many fundamental issues underlying selection of end points are substantial and in need of further empirical research.

These uncertainties make further methodological evaluation an important research priority. Some investigations are undertaken solely for methodological evaluation, but substantive studies assessing the relationships between exposure and disease can and should address these issues as well. Reports should include clear presentation of the rationale for the choice of reproductive measures and the underlying assumptions regarding such issues as the temporal relation with exposure and expectations regarding vulnerable subpopulations. Analyses should examine the relationship among multiple reproductive health measures to determine whether these measures are redundant and which seem to be the most sensitive indicators of reproductive risk. Environmental risk assessment in reproductive health would be well served by ambitious studies that creatively contend with these challenges.

The authors thank Sherry Selevan and Allen Wilcox for their helpful comments on the manuscript.

REFERENCES

- Chapman, R. M. Gonadal injury resulting from chemotherapy. *Am. J. Ind. Med.* 4: 149-161 (1983).
- Whorton, D., Kraus, R. M., Marshall, S., and Milby, T. H. Infertility in male pesticide workers. *Lancet* ii: 1259-1261 (1977).
- Baird, D. D., and Wilcox, A. J. Cigarette smoking associated with delayed conception. *J. Am. Med. Assoc.* 253: 2979-2983 (1985).
- Committee to Study the Prevention of Low Birth Weight. Preventing Low Birth Weight. National Academy of Sciences, Washington, DC, 1985.
- Wilson, J. G. Teratogenic effects of environmental chemicals. *Fed. Proc.* 36: 1698-1703 (1977).
- Barlow S. M., and Sullivan F. M. Reproductive Hazards of Industrial Chemicals. Academic Press, London, 1982.
- Office of Technology Assessment. Reproductive Health Hazards in the Workplace. U.S. Government Printing Office, Washington, DC, 1985.
- Baird, D. D., Wilcox, A. J., and Weinberg, C. R. Use of time to pregnancy to study environmental exposures. *Am. J. Epidemiol.* 124: 470-480 (1986).
- Wilcox, A. J. Surveillance of pregnancy loss in human populations. *Am. J. Ind. Med.* 4: 285-291 (1983).
- Vianna, N. J., and Polan, A. K. Incidence of low birth weight among Love Canal residents. *Science* 226: 1217-1219 (1984).
- Daling, J. R., Weiss, N. S., Metc, B. J., Chow, W. H., Soderstrom, R. M., Moore, D. E., Spadoni, L. R., and Stadel, B. V. Primary tubal infertility in relation to the use of an intrauterine device. *N. Engl. J. Med.* 312: 937-941 (1985).
- Shulte, P. A. Methodologic issues in the use of biologic markers in epidemiologic research. *Am. J. Epidemiol.* 126: 1006-1016 (1987).
- Whorton, M. D., and Foliart, D. E. Mutagenicity, carcinogenicity and reproductive effects of dibromochloropropane (DBCP). *Mutat. Res.* 123: 13-30 (1983).
- Rothman, K. J. Induction and latent periods. *Am. J. Epidemiol.* 114: 253-259 (1981).
- Selevan, S. G., and Lemasters G. K. The dose-response fallacy in human reproductive studies of toxic exposures. *J. Occup. Med.* 29: 451-454 (1987).
- Stein, Z., Susser, M., Warburton, D., Wittes, J., and Kline, J. Spontaneous abortion as a screening device. The effect of fetal survival on the incidence of birth defects. *Am. J. Epidemiol.* 102: 275-290 (1975).
- Buffler, P. A., and Aase, J. M. Genetic risks and environmental surveillance: epidemiological aspects of monitoring industrial populations for environmental mutagens. *J. Occup. Med.* 24: 305-314 (1982).
- Greenland, S., and Robins, J. M. Identifiability, exchangeability, and epidemiological confounding. *Int. J. Epidemiol.* 15: 413-419 (1986).
- Rothman, K. J., and Poole, C. A strengthening programme for weak associations. *Int. J. Epidemiol.* 17: 955-959 (1988).
- Kline, J., Stein, Z. A., Susser, M., and Warburton, D. Smoking: a risk factor for spontaneous abortion. *N. Engl. J. Med.* 297: 793-796 (1977).
- Selevan, S. G., Lindbohm, M. L., Hornung, R. W., and Hemminki, K. A study of occupational exposure to antineoplastic drugs and fetal loss in nurses. *N. Engl. J. Med.* 313: 1173-1178 (1985).
- Environmental Protection Agency. Proposed amendments to the guidelines for the health assessment of suspect developmental toxins. *Fed. Reg.* 54: 9386-9403 (1989).
- Rogan, W. J., Gladen, B. C., and Wilcox A. J. Potential reproductive and postnatal morbidity from exposure to polychlorinated biphenyls: epidemiologic considerations. *Environ. Health Perspect.* 60: 233-239 (1985).
- Nora, A. H., and Nora, J. J. A syndrome of multiple congenital anomalies associated with teratogenic exposure. *Arch. Environ. Health* 30: 17-21 (1975).
- Sterling, T. D., and Weinkam, J. I. Smoking characteristics by type of employment. *J. Occup. Med.* 18: 743-754 (1976).
- Weinburg G. B., Kuller, L. H., and Redmond, C. K. The relationship between the geographic distribution of lung cancer incidence and cigarette smoking in Allegheny County, Pennsylvania. *Am. J. Epidemiol.* 115: 40-58 (1982).
- Kline, J., and Stein Z. Spontaneous abortion (miscarriage). In: *Perinatal Epidemiology* (M. B. Bracken, Ed.), Oxford University Press, New York, 1984, pp. 23-51.
- Selevan, S. G. Design of pregnancy outcome studies of industrial exposures. In: *Occupational Hazards and Reproduction* (K. Hemminki, M. Sorsa, and H. Vainio, Eds.), Hemisphere Publishing Corporation, Washington, DC, 1985, pp. 219-229.
- Cates, W. J., Jr. Induced abortion. In: *Perinatal Epidemiology* (M. B. Bracken, Ed.), Oxford University Press, New York, 1984, pp. 301-324.
- Savitz, D. A., and Pearce, N. Control selection with incomplete case ascertainment. *Am. J. Epidemiol.* 127: 1109-1117 (1988).
- Mackeprang, M., Hay, S., and Lund, A. S. Completeness and accuracy of reporting of malformations on birth certificates. *Public Health Rep.* 87: 43-49 (1972).
- David, R. J. The quality and completeness of birthweight and gestational age data in computerized birth files. *Am. J. Public Health* 70: 964-973 (1980).
- Wilcox, A. J., and Horney, L. F. Accuracy of spontaneous abortion recall. *Am. J. Epidemiol.* 120: 727-733 (1984).
- Bean, J. A., Leper, J. D., Wallace, R. B., Sherman, B. M., and Jagger, H. Variations in the reporting of menstrual histories. *Am. J. Epidemiol.* 109: 181-185 (1979).